



Ein Spürhund für die Ausreißer in meteorologischen Daten

Informatik. Ein visuelles Analyse-Tool hilft dabei, Abweichungen und punktuelle Auffälligkeiten in großen Datensätzen ausfindig zu machen. Entwickelt für die digitalisierte Industrie, kann die Methode kombiniert mit statistisch-numerischen Verfahren auch Klimaprognosen verbessern.

VON CORNELIA GROBNER

Vorausschauende Wartung ist eine effiziente Strategie, um in Industriebetrieben Zeit und Kosten zu sparen. Voraussetzung dafür sind Sensoren, die bestimmte Vorgänge in Produktionsmaschinen, Temperaturen in Fertigungshallen, den Druck auf bestimmte Geräte und ähnliche Parameter in gewissen Zeitabständen messen. Anhand von minimalen Abweichungen können beginnende Abnutzungserscheinungen oder andere Probleme im Entstehen erkannt und behoben werden.

Die Informatikerin Johanna Schmidt entwickelt zusammen mit ihrem Team Tools, die die Auswertung und Interpretation dieser enormen Datenmengen erleichtern. Sie hat festgestellt, dass ihre Methoden auch bei der Analyse von Klima- und Energiedaten zum Einsatz kommen können. „Die Struktur der Daten ist ähnlich. Hier messen Sensoren ebenfalls in gewissen Zeitabständen Werte wie Temperatur oder Luftfeuchtigkeit, wodurch sich große - wir sagen dazu - Zeitserien ergeben“, erklärt sie. „Es war also naheliegend, unsere Datenvisualisierungstools auch dafür auszuprobieren.“

Schmidt leitet am Zentrum für Virtual Reality und Visualisierung (VRVis) in Wien

die Forschungsgruppe „Visual Analytics“. „Ich sehe es als meine Mission an, Menschen den Dialog mit Daten zu ermöglichen“, sagt sie. „Wir können visuelle Information einfach besser, schneller und teilweise auch genauer verarbeiten als Daten in Zahlen- oder Textform. Die visuelle Analyse eignet sich - gerade wenn es um komplexe Phänomene geht - als ideale Ergänzung zur statistischen oder numerischen Betrachtung von Daten.“

Historisches Mittel, reale Sonderfälle

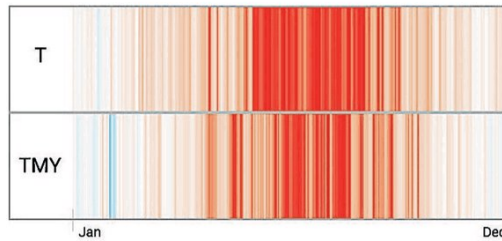
Und tatsächlich zeigte sich, dass sich die Visualisierungstools nicht nur bei Industrie-, sondern auch bei Klimadaten bewähren. Für den Testlauf analysierte die Informatikerin

gemeinsam mit ihrer Kollegin, der Bauingenieurwissenschaftlerin Milena Vuckovic, zwei verschiedene Datensätze. Zum einen handelte es sich um Wetterparameter von sogenannten typischen meteorologischen Jahren (TMY). Diese werden in der Bauplanung verwendet, um etwa die notwendige Kühlung für ein Gebäude zu berechnen oder den Heizaufwand abzuschätzen. Basis sind Zeitserien über ein Jahr mit stündlichen Werten zu Temperatur, Luftdruck, Windgeschwindigkeit und Windrichtung. „Die Datensätze werden aus der Kombination historischer Daten erstellt“, erklärt Schmidt. „Wir bereiteten unter anderem einen Datensatz von der Hohen Warte in

Wien von den vergangenen 18 Jahren visuell auf. Dazu kodierten wir die unterschiedlichen Temperaturen farblich.“ Ein bekanntes Problem von TMY-Daten ist, dass durch den Fokus auf die Mittelwerte, Extremsituationen verloren gehen. Deshalb nutzten die Forscherinnen für die Untersuchung zum anderen Realdaten des Jahres 2019.

„Im Vergleich der beiden visualisierten Datensätze sieht man, dass die realen Temperaturen viel höher waren als im typischen meteorologischen Jahr“, so Schmidt. Darüber hinaus offenbarte eine Analyse der Zeitserien zur Windrichtung weitere Diskrepanzen: „Der Südwind aus dem Modell Datensatz ist im Jahr 2019 verschwunden. Das weist möglicherweise auf eine Verschiebung durch den Klimawandel hin.“

Die im Fachmagazin *Data* publizierten Ergebnisse unterstreichen sowohl den zusätzlichen Nutzen einer visuellen Analyse als auch die Notwendigkeit, Modelldatensätze und reale Daten zu kombinieren. Schmidt: „Bisher genutzte, oft selbst gestrickte visuelle Lösungen sind limitiert, was große Datenmengen und die Skalierbarkeit angeht. Hier können wir helfen.“ So sind die am VRVis entwickelten Instrumente in der Lage, Datensätze mit Millionen von Datenpunkten flüssig zu verarbeiten.



Farbcodierte Diskrepanzen. Im gemittelten Datensatz (TMY) liegen die Höchsttemperaturen immer signifikant unter den realen Höchsttemperaturen (T).

[Foto: VRVis/Schmidt]